Poster: Feasibility of Malware Traffic Analysis through TLS-Encrypted Flow Visualization

IEEE International Conference on Network Protocols 2020
October 13-16, 2020

Dongeon Kim, Jihun Han, Jinwoo Lee, Heejun Roh
Korea University Sejong Campus, Sejong, Republic of Korea

Wonjun Lee
Korea University, Seoul, Republic of Korea
Motivation

Encrypted traffic across google

Network using TLS encryption is increasing

95% of traffic across google is encrypted

80% of enterprise traffic on the Zscaler cloud is encrypted

https://transparencyreport.google.com/https/overview?hl=en
Motivation

Nearly a quarter of malware now communicates using TLS

Motivation

- B. Anderson and D. McGrew, "Identifying encrypted malware traffic with contextual flow data," in *Proc. of AISec’16 (co-located with ACM CCS)*, Vienna, Austria, October 2016.

- Require fine-grained feature selection conducted by experts
- Need to conduct field-specific preprocessing for message field values
Our Proposal: TLS-Encrypted Flow Visualization

Image Format of TLS-Encrypted Flow
Our Proposal: TLS-Encrypted Flow Visualization

• TLS flow metadata have fruitful information to classify encrypted malware traffic

• Images can capture small changes yet retain the global message exchange pattern

• Different messages of a flow can be easily observed as a colored image
Images from Malware Families

Dridex  Gootkit  Hancitor  IcedID  Trickbot
Feasibility of Malware Traffic Analysis via Images

(a) Hancitor Sample Images  (b) Trickbot Sample Images
Experimental Results

Experimental Results

Resulting confusion matrices

93% Accuracy in Average

97% Accuracy in Average
Conclusion

• Malware using TLS will continue to increase in the future

• There needs to be new method to detect malware using TLS

• Both SVM and CNN had high accuracy, even though the images do not have similar patterns